P☉RTAL
USPTO

Search: ○ The ACM Digital Library   ◉ The Guide

SEARCH

THE GUIDE TO COMPUTING LITERATURE   Feedback

A performance evaluation of storing XML data in relational database management systems

Authors	Latifur Khan	University of Texas at Dallas, Richardson, TX
	Yan Rao	University of Texas at Dallas, Richardson, TX

Additional Information: abstract   references   index terms   collaborative colleagues   peer to peer

Tools and Actions: Review this Article

Save this Article to a Binder   Display Formats: BibTex   EndNote   ACM Ref

↑ ABSTRACT

XML is an emerging standard for the representation and exchange of Internet data. Along with document type definition (DTD), XML permits the execution of a collection of queries, using XPath to identify data in XML documents. In this paper we examine how XML data can be stored and queried using a standard relational database management system (RDBMS). For this, we propose a technique for automatic mapping from an XML document to relations within the RDBMS. We demonstrate that our novel approach preserves the nested structure of the XML

documents. By hiding database details we devise a seamless, transparent framework for user access to XML data. In order to achieve this, we propose a novel mechanism for translating an XPath query into an SQL statement. Furthermore, we propose efficient techniques for the construction of an XML document on the fly from the result set of the SQL statement. We also present findings in terms of query response time on the comparative performance of different techniques for the construction of an XML document on the fly.

⚘ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1. S. Abiteboul, D. Quass, J. MeHugh, J.Widom, J.Wiener. The Lorel Query Language for Semi-structured Data, International Journal on Digital Libraries, 1(1), pp. 68-88, April 1997

2. F. Bancilhon , G. Barbedette , V. Benzaken , C. Delobel , S. Gamerman , C. Lecluse , P. Pfeffer , P. Richard , F. Velez, The design and implementation of O2, Lecture notes in computer science on Advances in object-oriented database systems, p.1-22, September 1988, Bad Mu:9Ainster am Stein-Edernburg, Germany

3. T. Bray, J. Paoli, C.M. Sperberg-McQueen, Extensible Markup Language (XML) 1.0, http://www.w3.org/TR/REC-xml

4. Peter Buneman , Susan B. Davidson , Mary F. Fernandez , Dan Suciu, Adding Structure to Unstructured Data, Proceedings of the 6th International Conference on Database Theory, p.336-350, January 08-10, 1997

5. Alin Deutsch , Mary Fernandez , Daniela Florescu , Alon Levy , Dan Suciu, A query language for XML, Proceeding of the eighth international conference on World Wide Web, p.1155-1169, May 1999, Toronto, Canada

6. Alin Deutsch , Mary Fernandez , Dan Suciu, Storing semistructured data with STORED, Proceedings of the 1999 ACM SIGMOD international conference on Management of data, p.431-442, May 31-June 03, 1999, Philadelphia, Pennsylvania, United States

7. Mary Fernández , Daniela Florescu , Jaewoo Kang , Alon Levy , Dan Suciu, Catching the boat with Strudel: experiences with a Web-site management system, Proceedings of the 1998 ACM SIGMOD international conference on Management of data, p.414-425, June 01-04, 1998, Seattle, Washington, United States

8. D. Florescu, D. Kossman, A Performance Evaluation of Alternative Mapping Schemes for Storing XML Data in a Relational Database, Rapport de Recherche No. 3680 INRIA, Rocquencourt, France, May 1999

9 Jason McHugh , Serge Abiteboul , Roy Goldman , Dallas Quass , Jennifer Widom , Lore: a database management system for semistructured data, ACM SIGMOD Record, v.26 n.3, p.54-66, Sept. 1997 [doi> 10.1145/262762.262770]

10 Svetlozar Nestorov , Serge Abiteboul , Rajeev Motwani, Extracting schema from semistructured data, Proceedings of the 1998 ACM SIGMOD international conference on Management of data, p.295-306, June 01-04, 1998, Seattle, Washington, United States

11 Kurt A. Shoens , Allen Luniewski , Peter M. Schwarz , James W. Stamos , Joachim Thomas, II, The Rufus System: Information Organization for Semi-Structured Data, Proceedings of the 19th International Conference on Very Large Data Bases, p.97-107, August 24-27, 1993

12 Jayavel Shanmugasundaram , Kristin Tufte , Chun Zhang , Gang He , David J. DeWitt , Jeffrey F. Naughton, Relational Databases for Querying XML Documents: Limitations and Opportunities, Proceedings of the 25th International Conference on Very Large Data Bases, p.302-314, September 07-10, 1999

13 F. Tian, D. DeWitt, J. Chen, and C. Zhang, The Design and Performance Evaluation of Various XML Storage Strategies. Submitted for publication ,Computer Science, University of Wisconsin, Madison

14 R.V. Zwol, P. Apers, and A. Wilschut, Modeling and Querying Semi-structured Data With MOA, Workshop on Query Processing for Semi-structured Data and Non-standard Data Formats, 1999.

15 http://www.lotusnotes.com/, 1998.

16 XRel: a path-based approach to storage and retrieval of XML documents using relational databases. ACM Transactions on Internet Technology (TOIT), v.1 n.1, p.110-141, Aug. 2001 [doi> 10.1145/383034.383038]

17 Albrecht Schmidt , Martin L. Kersten , Menzo Windhouwer , Florian Waas, Efficient Relational Storage and Retrieval of XML Documents, Selected papers from the Third International Workshop WebDB 2000 on The World Wide Web and Databases, p.137-150, May 18-19, 2000

18 Mary F. Fernandez , Jérôme Siméon , Philip Wadler, An Algebra for XML Query, Proceedings of the 20th Conference on Foundations of Software Technology and Theoretical Computer Science, p.11-45, December 13-15, 2000

19 M. J. Carey and D. Florescu and Z. G. Ives and Y. Lu and J. Shanmugasundaram, E. J. Shekita and S. N. Subramanian",XPERANTO: Publishing Object-Relational Data as (XML)",WebDB (2000), pp. 105-110.

↑ INDEX TERMS

Primary Classification:

H. Information Systems

↳ H.3 INFORMATION STORAGE AND RETRIEVAL

↳ H.3.4 Systems and Software

↳ Subjects: Performance evaluation (efficiency and effectiveness)

Additional Classification:
H. Information Systems

↳ H.2 DATABASE MANAGEMENT

↳ H.2.4 Systems

↳ Subjects: Relational databases

I. Computing Methodologies

↳ I.7 DOCUMENT AND TEXT PROCESSING

↳ I.7.2 Document Preparation

↳ Nouns: XML

General Terms:
Design, Languages, Performance

Keywords:
DTD, SQL, XML, XPath, relational DBMS

↑ Collaborative Colleagues:

Latifur Khan:
colleagues

Yan Rao: colleagues

↑ Peer to Peer - Readers of this Article have also read:

- Constructing reality **Proceedings of the 11th annual international conference on Systems documentation**
  Douglas A. Powell , Norman R. Ball , Mansel W. Griffiths

- M³: a metamodel for data preprocessing **Proceedings of the 4th ACM international workshop on Data warehousing and OLAP**
  Anca Vaduva , Jörg-Uwe Kietz , Regina Zücker

- The effect of latency on user performance in Warcraft III **Proceedings of the 2nd workshop on Network and system support for games**
  Nathan Sheldon , Eric Girard , Seth Borg , Mark Claypool , Emmanuel Agu

- Learning subjective relevance to facilitate information access Proceedings of the fourth international conference on Information and knowledge management
James R. Chen , Nathalie Mathé

- Data structures for quadtree approximation and compression Communications of the ACM 28, 9
Hanan Samet

Useful downloads: Adobe Acrobat   QuickTime   Windows Media Player   Real Player